

## SEQUENCE LISTING

&lt;110&gt; CNRS

<120> STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE  
OF INTEREST, A METHOD OF OBTAINING THEM AND THEIR USES

&lt;130&gt; WOB 99 AB CNR AMYL

&lt;140&gt; US 09/980,771

&lt;141&gt; 2001-11-15

&lt;150&gt; FR 99/06494

&lt;151&gt; 1999-05-21

&lt;160&gt; 11

&lt;170&gt; PatentIn version 3.1

&lt;210&gt; 1

&lt;211&gt; 3117

&lt;212&gt; DNA

&lt;213&gt; Chlamydomonas reinhardtii

&lt;400&gt; 1

gccaggcgga	gcgtatggct	gttgccctcta	ccagccgccc	cagcagcgcg	cgtccctatcg	60
tcatcaatgc	cgcgtcggtc	ggtgtcaaga	agaccgcgaa	ccagctgctg	cgtgagcttg	120
ctcgtggctc	cgcacgcaag	tccacctcgc	gctcggtgtg	tactgggtgcc	actggtgcca	180
cttgccgctg	ggacatcggt	atggttgctg	ctgaggctgc	cccttgggtcc	aagacggggcg	240
gcctggggca	tgtgactggg	ggcctgccta	ttgagctggg	caagcgcggc	caccgcgtca	300
tgaccattgc	ccctcgctac	gaccagtagc	ctgacgcctg	ggacacctcg	gtggctcggtg	360
acatcatggg	cgagaagggt	cgtacttccc	actccatcaa	gaagggcggt	caccgcgtgt	420
ggattgacca	cccctgggtc	ctggccaagg	tctggggcaa	gaccggctcc	aagctgtacg	480
gcccgcgtc	cggcgctgac	tacctggaca	accacaagcg	cttcgccctg	ttctgcaagg	540
ccgctattga	ggctgcccgc	gtgctgccc	tgggcccccg	cgaggactgc	gtcttcgttg	600
ccaacgactg	gcactccgcc	ctgggtgccg	tcctgctgaa	ggacgagtag	cagcccaagg	660
gccagttcac	caaggccaag	tcgggtgctg	ctatccacaa	catcgccctc	cagggccgca	720
tgtgggagga	ggctttcaag	gacacgaagc	tgcgccaggc	cgcctttgac	aagctggcct	780
tctcggacgg	ctatgccaag	gtttacactg	aggccacccc	catggaggag	gacgagaagc	840
ccccgctgac	gggaaaagacc	tacaagaaga	tcaactgggt	gaaggggtgg	attatcgccg	900
ccgacaagct	ggtgactgtg	tgcgccaaact	acgcgaccga	gacgctgccc	gatgccgccc	960
gcgggtgtga	gctggacacc	gtcatccgcg	ccaagggcat	tgagggcatt	gtgaacggca	1020
tggacattga	ggagtggaa	cccaagaccg	acaagttcct	gtctgcgccc	tacgaccaga	1080
acagcgtcta	cgccggccaag	gccgcgcgca	aggaggccct	gcaggccgag	ctgggcccgc	1140
ctgtggaccc	caccgcccc	ctgttcgcct	tcctcgcccg	cctggaggag	cagaagggtg	1200
tggacatcat	cctggccgcc	ctgcccaaga	tcctggccac	ccccaaaggt	cagatcgcca	1260
tcctgggtac	cggcaaggcc	gcctacgaga	agctggtgaa	cgccatcggc	accaagtaca	1320
agggccgcgc	caagggcggt	gtcaagttct	cggcgccccc	ggcgccacat	ctcaccgccc	1380
gcgcgcgact	catgctgggt	ccctcgccgt	tcgagccctg	cggcctgac	cagctgcacg	1440
ccatgcacta	cggtagcgtg	cccgtggtag	cctccaccgg	cggcctgggt	gacaccgtca	1500
aggagggcgt	caccggcttc	cacatgggcg	ccctgaaccc	cgacaagctg	gacgaggctg	1560
acgcccagcg	cctggccgcc	accgtgcgcc	gtgccagcga	ggtgtttgcg	ggcgcccgct	1620
accccagat	ggtggccaac	tgcatacgcc	aggacctgtc	ctgggtccaag	cccggcccaga	1680
agtgggaggg	cctgctggag	gaggtggtgt	acggcaaggg	cggcgtggcc	accgccaaga	1740
aggaggagat	caaggtgccc	gttgccgaga	agatccccgg	cgacctgccc	gccgtgtcct	1800
acgcccccaa	caccctgaag	cccgtgtccg	cctccgtgga	gggcaacggc	gccgcgcgcg	1860
ccaaggtcgg	caccaccgcc	cccgcctatg	gcgcgtggcg	cgcgaccacc	ccctcggggc	1920
cctcgccccg	cgccgcacc	cccaaggtga	ccacctacaa	gcccgcctg	cccggccaccg	1980
ccaagcccaa	gaccgctggc	ctcaagctgg	ccggtgaggg	ctccaccacc	tcgacctcgg	2040
agaacggcgc	tgccctccaa	ggcaacggca	acggtgcctc	ggcctccaa	acctcggtcg	2100
ccaagccccc	ggtctccgcc	gccacccgca	agtccgccta	aagcggcagt	agccgcagag	2160

```

gctggcgacag catgagcggc tcgaccaaag ctgtggcagg aacggctgta gcagcggcag 2220
gcggccgcca ccggcgagga gcaggcttgc ggcagcgagg gcgatgagct tagcggccgt 2280
gagcatggca ggcggaaacg tgtgtactga aatgtggtgc atgagagtgt cgtgctgtaa 2340
tgaagtcggt tttgcgagac cggagaaacg ccggtttggt tttgtagtgc agggcctgtg 2400
gtttcggttt tgcccaagtc caaaagaaga gtaacgaaac tgtagcagta gcagagcact 2460
tgcgcgggcg gcgaccacg ccggcccgtg cgcagcctgt cctgccctca gccttgtgat 2520
tcggcgggcaa gagggcgggt ctgtacactc catccattcc aggatttttg caggctgcct 2580
gagagtttgc cattttgttg gacgtgagcg gcgggacggc cgcgcccggc tctcctaccg 2640
cctccggcaa cggagaagtg ggaggcgtg tagcccgggtg acccccctaat gtagaggatg 2700
ggatacataa gagcgtgttg aatggtggtg aaagaggagg ggcctgggtc gccctcgat 2760
ggttttgttg aggtgcagac ggcaccgtcg gcgtcaaagg ccctcgcaag gcccggtgc 2820
cttgggctca tttttggtgc ccgtcgatga tgagagattg gccagcgggt ttttgaggct 2880
ggctcgaagc gagggtttgt ggaagtggag cgaggagggt tggagaaaga ggcggacatg 2940
cttgactgga ggtacacaaa gtggagcgtg cgacggcacg gaggcattgg cggactattg 3000
accagtagt gtggaaagta gttggacctg aattctttga gagtaccgcg cattaatccg 3060
tgagagagta acaaagatgg cacctgaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa 3117

```

```

<210> 2
<211> 2124
<212> DNA
<213> artificial sequence

```

```

<220>
<223> fragment of the complete sequence of cDNA coding
      for the GBSSI of Chlamydomonas reinhardtii

```

```

<220>
<221> CDS
<222> (1)..(2124)
<223>

```

```

<400> 2
atg gct gtt gcc tct acc agc cgc ccc agc agc gcg cgt cct atc gtc 48
Met Ala Val Ala Ser Thr Ser Arg Pro Ser Ser Ala Arg Pro Ile Val
1 5 10 15

atc aat gcc gcg tcg ttc ggt gtc aag aag acc gcg aac cag ctg ctg 96
Ile Asn Ala Ala Ser Phe Gly Val Lys Lys Thr Ala Asn Gln Leu Leu
20 25 30

cgt gag ctt gct cgt ggc tcc gca cgc aag tcc acc tcg cgc tcg gct 144
Arg Glu Leu Ala Arg Gly Ser Ala Arg Lys Ser Thr Ser Arg Ser Ala
35 40 45

gtt act ggt gcc act ggt gcc act tgc gcg ctg gac atc gtg atg gtt 192
Val Thr Gly Ala Thr Gly Ala Thr Cys Ala Leu Asp Ile Val Met Val
50 55 60

gct gct gag gtc gcc cct tgg tcc aag acg ggc ggc ctg ggc gat gtg 240
Ala Ala Glu Val Ala Pro Trp Ser Lys Thr Gly Gly Leu Gly Asp Val
65 70 75 80

act ggt ggc ctg cct att gag ctg gtc aag cgc ggc cac cgc gtc atg 288
Thr Gly Gly Leu Pro Ile Glu Leu Val Lys Arg Gly His Arg Val Met
85 90 95

acc att gcc cct cgc tac gac cag tac gct gac gcc tgg gac acc tcg 336
Thr Ile Ala Pro Arg Tyr Asp Gln Tyr Ala Asp Ala Trp Asp Thr Ser
100 105 110

```

gtg gtc gtg gac atc atg ggc gag aag gtc cgc tac ttc cac tcc atc Val Val Val Asp Ile Met Gly Glu Lys Val Arg Tyr Phe His Ser Ile 115 120 125	384
aag aag ggc gtg cac cgc gtg tgg att gac cac ccc tgg ttc ctg gcc Lys Lys Gly Val His Arg Val Trp Ile Asp His Pro Trp Phe Leu Ala 130 135 140	432
aag gtc tgg ggc aag acc ggc tcc aag ctg tac ggc ccc cgc tcc ggc Lys Val Trp Gly Lys Thr Gly Ser Lys Leu Tyr Gly Pro Arg Ser Gly 145 150 155 160	480
gct gac tac ctg gac aac cac aag cgc ttc gcc ctg ttc tgc aag gcc Ala Asp Tyr Leu Asp Asn His Lys Arg Phe Ala Leu Phe Cys Lys Ala 165 170 175	528
gct att gag gct gcc cgc gtg ctg ccc ttc ggc ccc ggc gag gac tgc Ala Ile Glu Ala Ala Arg Val Leu Pro Phe Gly Pro Gly Glu Asp Cys 180 185 190	576
gtc ttc gtg gcc aac gac tgg cac tcc gcc ctg gtg ccc gtc ctg ctg Val Phe Val Ala Asn Asp Trp His Ser Ala Leu Val Pro Val Leu Leu 195 200 205	624
aag gac gag tac cag ccc aag ggc cag ttc acc aag gcc aag tcg gtg Lys Asp Glu Tyr Gln Pro Lys Gly Gln Phe Thr Lys Ala Lys Ser Val 210 215 220	672
ctg gct atc cac aac atc gcc ttc cag ggc cgc atg tgg gag gag gct Leu Ala Ile His Asn Ile Ala Phe Gln Gly Arg Met Trp Glu Glu Ala 225 230 235 240	720
ttc aag gac acg aag ctg ccc cag gcc gcc ttt gac aag ctg gcc ttc Phe Lys Asp Thr Lys Leu Pro Gln Ala Ala Phe Asp Lys Leu Ala Phe 245 250 255	768
tcg gac ggc tat gcc aag gtt tac act gag gcc acc ccc atg gag gag Ser Asp Gly Tyr Ala Lys Val Tyr Thr Glu Ala Thr Pro Met Glu Glu 260 265 270	816
gac gag aag ccc ccg ctg acg gga aag acc tac aag aag atc aac tgg Asp Glu Lys Pro Pro Leu Thr Gly Lys Thr Tyr Lys Lys Ile Asn Trp 275 280 285	864
ctg aag ggt ggc att atc gcc gcc gac aag ctg gtg act gtg tcg ccc Leu Lys Gly Gly Ile Ile Ala Ala Asp Lys Leu Val Thr Val Ser Pro 290 295 300	912
aac tac gcg acc gag atc gct gcc gat gcc gcc ggc ggt gtg gag ctg Asn Tyr Ala Thr Glu Ile Ala Ala Asp Ala Ala Gly Gly Val Glu Leu 305 310 315 320	960
gac acc gtc atc cgc gcc aag ggc att gag ggc att gtg aac ggc atg Asp Thr Val Ile Arg Ala Lys Gly Ile Glu Gly Ile Val Asn Gly Met 325 330 335	1008
gac att gag gag tgg aac ccc aag acc gac aag ttc ctg tct gcg ccc Asp Ile Glu Glu Trp Asn Pro Lys Thr Asp Lys Phe Leu Ser Ala Pro 340 345 350	1056

tac gac cag aac agc gtc tac gcc ggc aag gcc gcc gcc aag gag gcc Tyr Asp Gln Asn Ser Val Tyr Ala Gly Lys Ala Ala Ala Lys Glu Ala 355 360 365	1104
ctg cag gcc gag ctg ggc ctg cct gtg gac ccc acc gcc ccc ctg ttc Leu Gln Ala Glu Leu Gly Leu Pro Val Asp Pro Thr Ala Pro Leu Phe 370 375 380	1152
gcc ttc atc ggc cgc ctg gag gag cag aag ggt gtg gac atc atc ctg Ala Phe Ile Gly Arg Leu Glu Glu Gln Lys Gly Val Asp Ile Ile Leu 385 390 395 400	1200
gcc gcc ctg ccc aag atc ctg gcc acc ccc aag gtg cag atc gcc atc Ala Ala Leu Pro Lys Ile Leu Ala Thr Pro Lys Val Gln Ile Ala Ile 405 410 415	1248
ctg ggt acc ggc aag gcc gcc tac gag aag ctg gtg aac gcc atc ggc Leu Gly Thr Gly Lys Ala Ala Tyr Glu Lys Leu Val Asn Ala Ile Gly 420 425 430	1296
acc aag tac aag ggc cgc gcc aag ggc gtg gtc aag ttc tcg gcg ccc Thr Lys Tyr Lys Gly Arg Ala Lys Gly Val Val Lys Phe Ser Ala Pro 435 440 445	1344
ctg gcg cac atg ctc acc gcc ggc gcc gac ttc atg ctg gtg ccc tcg Leu Ala His Met Leu Thr Ala Gly Ala Asp Phe Met Leu Val Pro Ser 450 455 460	1392
cgc ttc gag ccc tgc ggc ctg atc cag ctg cac gcc atg cac tac ggt Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu His Ala Met His Tyr Gly 465 470 475 480	1440
acc gtg ccc gtg gta gcc tcc acc ggc ggc ctg gtc gac acc gtc aag Thr Val Pro Val Val Ala Ser Thr Gly Gly Leu Val Asp Thr Val Lys 485 490 495	1488
gag ggc gtc acc ggc ttc cac atg ggc gcc ctg aac ccc gac aag ctg Glu Gly Val Thr Gly Phe His Met Gly Ala Leu Asn Pro Asp Lys Leu 500 505 510	1536
gac gag gct gac gcc gac gcc ctg gcc gcc acc gtg cgc cgt gcc agc Asp Glu Ala Asp Ala Asp Ala Leu Ala Ala Thr Val Arg Arg Ala Ser 515 520 525	1584
gag gtg ttt gcg ggc ggc cgc tac ccc gag atg gtg gcc aac tgc atc Glu Val Phe Ala Gly Gly Arg Tyr Pro Glu Met Val Ala Asn Cys Ile 530 535 540	1632
agc cag gac ctg tcc tgg tcc aag ccc gcc cag aag tgg gag ggc ctg Ser Gln Asp Leu Ser Trp Ser Lys Pro Ala Gln Lys Trp Glu Gly Leu 545 550 555 560	1680
ctg gag gag gtg gtg tac ggc aag ggc ggc gtg gcc acc gcc aag aag Leu Glu Glu Val Val Tyr Gly Lys Gly Gly Val Ala Thr Ala Lys Lys 565 570 575	1728
gag gag atc aag gtg ccc gtt gcc gag aag atc ccc ggc gac ctg ccc Glu Glu Ile Lys Val Pro Val Ala Glu Lys Ile Pro Gly Asp Leu Pro 580 585 590	1776

gdc gtg tcc tac gcc ccc aac acc ctg aag ccc gtg tcc gcc tcc gtg 1824  
Ala Val Ser Tyr Ala Pro Asn Thr Leu Lys Pro Val Ser Ala Ser Val  
595 600 605

gag ggc aac ggc gcc gcc gcg ccc aag gtc ggc acc acc gcc ccc gcc 1872  
Glu Gly Asn Gly Ala Ala Ala Pro Lys Val Gly Thr Thr Ala Pro Ala  
610 615 620

atg ggc gcg tgg cgc gcg acc acc ccc tgg ggc ccc tgg ccc gcc gcc 1920  
Met Gly Ala Trp Arg Ala Thr Thr Pro Ser Gly Pro Ser Pro Ala Ala  
625 630 635 640

gcc acc ccc aag gtg acc acc tac aag ccc gcc ctg ccc gcc acc gcc 1968  
Ala Thr Pro Lys Val Thr Thr Tyr Lys Pro Ala Leu Pro Ala Thr Ala  
645 650 655

aag ccc aag acc gct ggc ctc aag ctg gcc ggt gag gcc tcc acc acc 2016  
Lys Pro Lys Thr Ala Gly Leu Lys Leu Ala Gly Glu Ala Ser Thr Thr  
660 665 670

tcg acc tcg gag aac ggc gct gcc tcc aac ggc aac ggc aac ggt gcc 2064  
Ser Thr Ser Glu Asn Gly Ala Ala Ser Asn Gly Asn Gly Asn Gly Ala  
675 680 685

tcg gcc tcc aag acc tcg gct gcc aag ccc ctg gtc tcc gcc gcc acc 2112  
Ser Ala Ser Lys Thr Ser Ala Ala Lys Pro Leu Val Ser Ala Ala Thr  
690 695 700

cgc aag tcc gcc 2124  
Arg Lys Ser Ala  
705

 $\langle 210 \rangle$  3

<211> 708

&lt;212&gt; PRT

<213> artificial sequence

 $\langle 220 \rangle$ 

<223> fragment of the complete sequence of cDNA coding for the GBSSI of *Chlamydomonas reinhardtii*

<400> 3

Met Ala Val Ala Ser Thr Ser Arg Pro Ser Ser Ala Arg Pro Ile Val  
1 5 10 15

Ile Asn Ala Ala Ser Phe Gly Val Lys Lys Thr Ala Asn Gln Leu Leu  
20 25 30

Arg Glu Leu Ala Arg Gly Ser Ala Arg Lys Ser Thr Ser Arg Ser Ala  
35 40 45

Val Thr Gly Ala Thr Gly Ala Thr Cys Ala Leu Asp Ile Val Met Val  
50 55 60

Ala Ala Glu Val Ala Pro Trp Ser Lys Thr Gly Gly Leu Gly Asp Val  
65 70 75 80

Thr Gly Gly Leu Pro Ile Glu.Leu Val Lys Arg Gly His Arg Val Met  
85 90 95

Thr	Ile	Ala	Pro	Arg	Tyr	Asp	Gln	Tyr	Ala	Asp	Ala	Trp	Asp	Thr	Ser		
			100					105					110				
Val	Val	Val	Asp	Ile	Met	Gly	Glu	Lys	Val	Arg	Tyr	Phe	His	Ser	Ile		
		115					120					125					
Lys	Lys	Gly	Val	His	Arg	Val	Trp	Ile	Asp	His	Pro	Trp	Phe	Leu	Ala		
	130					135					140						
Lys	Val	Trp	Gly	Lys	Thr	Gly	Ser	Lys	Leu	Tyr	Gly	Pro	Arg	Ser	Gly		
145					150					155					160		
Ala	Asp	Tyr	Leu	Asp	Asn	His	Lys	Arg	Phe	Ala	Leu	Phe	Cys	Lys	Ala		
				165					170					175			
Ala	Ile	Glu	Ala	Ala	Arg	Val	Leu	Pro	Phe	Gly	Pro	Gly	Glu	Asp	Cys		
			180					185					190				
Val	Phe	Val	Ala	Asn	Asp	Trp	His	Ser	Ala	Leu	Val	Pro	Val	Leu	Leu		
		195					200					205					
Lys	Asp	Glu	Tyr	Gln	Pro	Lys	Gly	Gln	Phe	Thr	Lys	Ala	Lys	Ser	Val		
	210					215					220						
Leu	Ala	Ile	His	Asn	Ile	Ala	Phe	Gln	Gly	Arg	Met	Trp	Glu	Glu	Ala		
225					230				235						240		
Phe	Lys	Asp	Thr	Lys	Leu	Pro	Gln	Ala	Ala	Phe	Asp	Lys	Leu	Ala	Phe		
				245					250					255			
Ser	Asp	Gly	Tyr	Ala	Lys	Val	Tyr	Thr	Glu	Ala	Thr	Pro	Met	Glu	Glu		
			260					265					270				
Asp	Glu	Lys	Pro	Pro	Leu	Thr	Gly	Lys	Thr	Tyr	Lys	Lys	Ile	Asn	Trp		
	275						280					285					
Leu	Lys	Gly	Gly	Ile	Ile	Ala	Ala	Asp	Lys	Leu	Val	Thr	Val	Ser	Pro		
	290					295					300						
Asn	Tyr	Ala	Thr	Glu	Ile	Ala	Ala	Asp	Ala	Ala	Gly	Gly	Val	Glu	Leu		
305					310				315					320			
Asp	Thr	Val	Ile	Arg	Ala	Lys	Gly	Ile	Glu	Gly	Ile	Val	Asn	Gly	Met		
				325					330					335			
Asp	Ile	Glu	Glu	Trp	Asn	Pro	Lys	Thr	Asp	Lys	Phe	Leu	Ser	Ala	Pro		
			340					345					350				
Tyr	Asp	Gln	Asn	Ser	Val	Tyr	Ala	Gly	Lys	Ala	Ala	Ala	Lys	Glu	Ala		
		355					360					365					
Leu	Gln	Ala	Glu	Leu	Gly	Leu	Pro	Val	Asp	Pro	Thr	Ala	Pro	Leu	Phe		
	370					375					380						
Ala	Phe	Ile	Gly	Arg	Leu	Glu	Glu	Gln	Lys	Gly	Val	Asp	Ile	Ile	Leu		
385					390					395					400		
Ala	Ala	Leu	Pro	Lys	Ile	Leu	Ala	Thr	Pro	Lys	Val	Gln	Ile	Ala	Ile		
				405					410					415			

Leu Gly Thr Gly Lys Ala Ala Tyr Glu Lys Leu Val Asn Ala Ile Gly  
 420 425 430  
 Thr Lys Tyr Lys Gly Arg Ala Lys Gly Val Val Lys Phe Ser Ala Pro  
 435 440 445  
 Leu Ala His Met Leu Thr Ala Gly Ala Asp Phe Met Leu Val Pro Ser  
 450 455 460  
 Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu His Ala Met His Tyr Gly  
 465 470 475 480  
 Thr Val Pro Val Val Ala Ser Thr Gly Gly Leu Val Asp Thr Val Lys  
 485 490 495  
 Glu Gly Val Thr Gly Phe His Met Gly Ala Leu Asn Pro Asp Lys Leu  
 500 505 510  
 Asp Glu Ala Asp Ala Asp Ala Leu Ala Ala Thr Val Arg Arg Ala Ser  
 515 520 525  
 Glu Val Phe Ala Gly Gly Arg Tyr Pro Glu Met Val Ala Asn Cys Ile  
 530 535 540  
 Ser Gln Asp Leu Ser Trp Ser Lys Pro Ala Gln Lys Trp Glu Gly Leu  
 545 550 555 560  
 Leu Glu Glu Val Val Tyr Gly Lys Gly Gly Val Ala Thr Ala Lys Lys  
 565 570 575  
 Glu Glu Ile Lys Val Pro Val Ala Glu Lys Ile Pro Gly Asp Leu Pro  
 580 585 590  
 Ala Val Ser Tyr Ala Pro Asn Thr Leu Lys Pro Val Ser Ala Ser Val  
 595 600 605  
 Glu Gly Asn Gly Ala Ala Ala Pro Lys Val Gly Thr Thr Ala Pro Ala  
 610 615 620  
 Met Gly Ala Trp Arg Ala Thr Thr Pro Ser Gly Pro Ser Pro Ala Ala  
 625 630 635 640  
 Ala Thr Pro Lys Val Thr Thr Tyr Lys Pro Ala Leu Pro Ala Thr Ala  
 645 650 655  
 Lys Pro Lys Thr Ala Gly Leu Lys Leu Ala Gly Glu Ala Ser Thr Thr  
 660 665 670  
 Ser Thr Ser Glu Asn Gly Ala Ala Ser Asn Gly Asn Gly Asn Gly Ala  
 675 680 685  
 Ser Ala Ser Lys Thr Ser Ala Ala Lys Pro Leu Val Ser Ala Ala Thr  
 690 695 700  
 Arg Lys Ser Ala  
 705

<210> 4  
 <211> 1953  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> fragment of the complete sequence of cDNA coding  
 for the GBSSI of *Chlamydomonas reinhardtii* and  
 coding for the mature GBSSI protein

<220>  
 <221> CDS  
 <222> (1)..(1953)  
 <223>

<400> 4  
 gcg ctg gac atc gtg atg gtt gct gct gag gtc gcc cct tgg tcc aag 48  
 Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys  
 1 5 10 15  
 acg ggc ggc ctg ggc gat gtg act ggt ggc ctg cct att gag ctg gtc 96  
 Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val  
 20 25 30  
 aag cgc ggc cac cgc gtc atg acc att gcc cct cgc tac gac cag tac 144  
 Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr  
 35 40 45  
 gct gac gcc tgg gac acc tcg gtg gtc gtg gac atc atg ggc gag aag 192  
 Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys  
 50 55 60  
 gtc cgc tac ttc cac tcc atc aag aag ggc gtg cac cgc gtg tgg att 240  
 Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile  
 65 70 75 80  
 gac cac ccc tgg ttc ctg gcc aag gtc tgg ggc aag acc ggc tcc aag 288  
 Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys  
 85 90 95  
 ctg tac ggc ccc cgc tcc ggc gct gac tac ctg gac aac cac aag cgc 336  
 Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg  
 100 105 110  
 ttc gcc ctg ttc tgc aag gcc gct att gag gct gcc cgc gtg ctg ccc 384  
 Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro  
 115 120 125  
 ttc ggc ccc ggc gag gac tgc gtc ttc gtg gcc aac gac tgg cac tcc 432  
 Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser  
 130 135 140  
 gcc ctg gtg ccc gtc ctg ctg aag gac gag tac cag ccc aag ggc cag 480  
 Ala Leu Val Pro Val Leu Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln  
 145 150 155 160  
 ttc acc aag gcc aag tcg gtg ctg gct atc cac aac atc gcc ttc cag 528  
 Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln  
 165 170 175



ggc cgc atg tgg gag gag gct ttc aag gac acg aag ctg ccc cag gcc	576
Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala	
180 185 190	
gcc ttt gac aag ctg gcc ttc tcg gac ggc tat gcc aag gtt tac act	624
Ala Phe Asp Lys Leu Ala Phe Ser Asp Gly Tyr Ala Lys Val Tyr Thr	
195 200 205	
gag gcc acc ccc atg gag gag gac gag aag ccc ccg ctg acg gga aag	672
Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys	
210 215 220	
acc tac aag aag atc aac tgg ctg aag ggt ggc att atc gcc gcc gac	720
Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp	
225 230 235 240	
aag ctg gtg act gtg tcg ccc aac tac gcg acc gag atc gct gcc gat	768
Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp	
245 250 255	
gcc gcc ggc ggt gtg gag ctg gac acc gtc atc cgc gcc aag ggc att	816
Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile	
260 265 270	
gag ggc att gtg aac ggc atg gac att gag gag tgg aac ccc aag acc	864
Glu Gly Ile Val Asn Gly Met Asp Ile Glu Glu Trp Asn Pro Lys Thr	
275 280 285	
gac aag ttc ctg tct gcg ccc tac gac cag aac agc gtc tac gcc ggc	912
Asp Lys Phe Leu Ser Ala Pro Tyr Asp Gln Asn Ser Val Tyr Ala Gly	
290 295 300	
aag gcc gcc gcc aag gag gcc ctg cag gcc gag ctg ggc ctg cct gtg	960
Lys Ala Ala Ala Lys Glu Ala Leu Gln Ala Glu Leu Gly Leu Pro Val	
305 310 315 320	
gac ccc acc gcc ccc ctg ttc gcc ttc atc ggc cgc ctg gag gag cag	1008
Asp Pro Thr Ala Pro Leu Phe Ala Phe Ile Gly Arg Leu Glu Glu Gln	
325 330 335	
aag ggt gtg gac atc atc ctg gcc gcc ctg ccc aag atc ctg gcc acc	1056
Lys Gly Val Asp Ile Ile Leu Ala Ala Leu Pro Lys Ile Leu Ala Thr	
340 345 350	
ccc aag gtg cag atc gcc atc ctg ggt acc ggc aag gcc gcc tac gag	1104
Pro Lys Val Gln Ile Ala Ile Leu Gly Thr Gly Lys Ala Ala Tyr Glu	
355 360 365	
aag ctg gtg aac gcc atc ggc acc aag tac aag ggc cgc gcc aag ggc	1152
Lys Leu Val Asn Ala Ile Gly Thr Lys Tyr Lys Gly Arg Ala Lys Gly	
370 375 380	
gtg gtc aag ttc tcg gcg ccc ctg gcg cac atg ctc acc gcc ggc gcc	1200
Val Val Lys Phe Ser Ala Pro Leu Ala His Met Leu Thr Ala Gly Ala	
385 390 395 400	
gac ttc atg ctg gtg ccc tcg cgc ttc gag ccc tgc ggc ctg atc cag	1248
Asp Phe Met Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln	
405 410 415	

ctg cac gcc atg cac tac ggt acc gtg ccc gtg gta gcc tcc acc ggc Leu His Ala Met His Tyr Gly Thr Val Pro Val Val Ala Ser Thr Gly 420 425 430	1296
ggc ctg gtc gac acc gtc aag gag ggc gtc acc ggc ttc cac atg ggc Gly Leu Val Asp Thr Val Lys Glu Gly Val Thr Gly Phe His Met Gly 435 440 445	1344
gcc ctg aac ccc gac aag ctg gac gag gct gac gcc gac gcc ctg gcc Ala Leu Asn Pro Asp Lys Leu Asp Glu Ala Asp Ala Asp Ala Leu Ala 450 455 460	1392
gcc acc gtg cgc cgt gcc agc gag gtg ttt gcg ggc ggc cgc tac ccc Ala Thr Val Arg Arg Ala Ser Glu Val Phe Ala Gly Gly Arg Tyr Pro 465 470 475 480	1440
gag atg gtg gcc aac tgc atc agc cag gac ctg tcc tgg tcc aag ccc Glu Met Val Ala Asn Cys Ile Ser Gln Asp Leu Ser Trp Ser Lys Pro 485 490 495	1488
gcc cag aag tgg gag ggc ctg ctg gag gag gtg gtg tac ggc aag ggc Ala Gln Lys Trp Glu Gly Leu Leu Glu Glu Val Val Tyr Gly Lys Gly 500 505 510	1536
ggc gtg gcc acc gcc aag aag gag gag atc aag gtg ccc gtt gcc gag Gly Val Ala Thr Ala Lys Lys Glu Glu Ile Lys Val Pro Val Ala Glu 515 520 525	1584
aag atc ccc ggc gac ctg ccc gcc gtg tcc tac gcc ccc aac acc ctg Lys Ile Pro Gly Asp Leu Pro Ala Val Ser Tyr Ala Pro Asn Thr Leu 530 535 540	1632
aag ccc gtg tcc gcc tcc gtg gag ggc aac ggc gcc gcc gcg ccc aag Lys Pro Val Ser Ala Ser Val Glu Gly Asn Gly Ala Ala Ala Pro Lys 545 550 555 560	1680
gtc ggc acc acc gcc ccc gcc atg ggc gcg tgg cgc gcg acc acc ccc Val Gly Thr Thr Ala Pro Ala Met Gly Ala Trp Arg Ala Thr Thr Pro 565 570 575	1728
tcg ggc ccc tcg ccc gcc gcc gcc acc ccc aag gtg acc acc tac aag Ser Gly Pro Ser Pro Ala Ala Ala Thr Pro Lys Val Thr Thr Tyr Lys 580 585 590	1776
ccc gcc ctg ccc gcc acc gcc aag ccc aag acc gct ggc ctc aag ctg Pro Ala Leu Pro Ala Thr Ala Lys Pro Lys Thr Ala Gly Leu Lys Leu 595 600 605	1824
gcc ggt gag gcc tcc acc acc tcg acc tcg gag aac ggc gct gcc tcc Ala Gly Glu Ala Ser Thr Thr Ser Thr Ser Glu Asn Gly Ala Ala Ser 610 615 620	1872
aac ggc aac ggc aac ggt gcc tcg gcc tcc aag acc tcg gct gcc aag Asn Gly Asn Gly Asn Gly Ala Ser Ala Ser Lys Thr Ser Ala Ala Lys 625 630 635 640	1920
ccc ctg gtc tcc gcc gcc acc cgc aag tcc gcc Pro Leu Val Ser Ala Ala Thr Arg Lys Ser Ala 645 650	1953

<210> 5  
 <211> 651  
 <212> PRT  
 <213> artificial sequence

<220>  
 <223> fragment of the complete sequence of cDNA coding  
 for the GBSSI of *Chlamydomonas reinhardtii* and  
 coding for the mature GBSSI protein

<400> 5  
 Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys  
 1 5 10 15  
 Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val  
 20 25 30  
 Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr  
 35 40 45  
 Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys  
 50 55 60  
 Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile  
 65 70 75 80  
 Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys  
 85 90 95  
 Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg  
 100 105 110  
 Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro  
 115 120 125  
 Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser  
 130 135 140  
 Ala Leu Val Pro Val Leu Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln  
 145 150 155 160  
 Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln  
 165 170 175  
 Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala  
 180 185 190  
 Ala Phe Asp Lys Leu Ala Phe Ser Asp Gly Tyr Ala Lys Val Tyr Thr  
 195 200 205  
 Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys  
 210 215 220  
 Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp  
 225 230 235 240  
 Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp  
 245 250 255  
 Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile  
 260 265 270

Glu	Gly	Ile	Val	Asn	Gly	Met	Asp	Ile	Glu	Glu	Trp	Asn	Pro	Lys	Thr	275	280	285
Asp	Lys	Phe	Leu	Ser	Ala	Pro	Tyr	Asp	Gln	Asn	Ser	Val	Tyr	Ala	Gly	290	295	300
Lys	Ala	Ala	Ala	Lys	Glu	Ala	Leu	Gln	Ala	Glu	Leu	Gly	Leu	Pro	Val	305	310	315
Asp	Pro	Thr	Ala	Pro	Leu	Phe	Ala	Phe	Ile	Gly	Arg	Leu	Glu	Glu	Gln	325	330	335
Lys	Gly	Val	Asp	Ile	Ile	Leu	Ala	Ala	Leu	Pro	Lys	Ile	Leu	Ala	Thr	340	345	350
Pro	Lys	Val	Gln	Ile	Ala	Ile	Leu	Gly	Thr	Gly	Lys	Ala	Ala	Tyr	Glu	355	360	365
Lys	Leu	Val	Asn	Ala	Ile	Gly	Thr	Lys	Tyr	Lys	Gly	Arg	Ala	Lys	Gly	370	375	380
Val	Val	Lys	Phe	Ser	Ala	Pro	Leu	Ala	His	Met	Leu	Thr	Ala	Gly	Ala	385	390	395
Asp	Phe	Met	Leu	Val	Pro	Ser	Arg	Phe	Glu	Pro	Cys	Gly	Leu	Ile	Gln	405	410	415
Leu	His	Ala	Met	His	Tyr	Gly	Thr	Val	Pro	Val	Val	Ala	Ser	Thr	Gly	420	425	430
Gly	Leu	Val	Asp	Thr	Val	Lys	Glu	Gly	Val	Thr	Gly	Phe	His	Met	Gly	435	440	445
Ala	Leu	Asn	Pro	Asp	Lys	Leu	Asp	Glu	Ala	Asp	Ala	Asp	Ala	Leu	Ala	450	455	460
Ala	Thr	Val	Arg	Arg	Ala	Ser	Glu	Val	Phe	Ala	Gly	Gly	Arg	Tyr	Pro	465	470	475
Glu	Met	Val	Ala	Asn	Cys	Ile	Ser	Gln	Asp	Leu	Ser	Trp	Ser	Lys	Pro	485	490	495
Ala	Gln	Lys	Trp	Glu	Gly	Leu	Leu	Glu	Glu	Val	Val	Tyr	Gly	Lys	Gly	500	505	510
Gly	Val	Ala	Thr	Ala	Lys	Lys	Glu	Glu	Ile	Lys	Val	Pro	Val	Ala	Glu	515	520	525
Lys	Ile	Pro	Gly	Asp	Leu	Pro	Ala	Val	Ser	Tyr	Ala	Pro	Asn	Thr	Leu	530	535	540
Lys	Pro	Val	Ser	Ala	Ser	Val	Glu	Gly	Asn	Gly	Ala	Ala	Ala	Pro	Lys	545	550	555
Val	Gly	Thr	Thr	Ala	Pro	Ala	Met	Gly	Ala	Trp	Arg	Ala	Thr	Thr	Pro	565	570	575
Ser	Gly	Pro	Ser	Pro	Ala	Ala	Ala	Thr	Pro	Lys	Val	Thr	Thr	Tyr	Lys	580	585	590

Pro Ala Leu Pro Ala Thr Ala Lys Pro Lys Thr Ala Gly Leu Lys Leu  
595 600 605

Ala Gly Glu Ala Ser Thr Thr Ser Thr Ser Glu Asn Gly Ala Ala Ser  
610 615 620

Asn Gly Asn Gly Asn Gly Ala Ser Ala Ser Lys Thr Ser Ala Ala Lys  
625 630 635 640

Pro Leu Val Ser Ala Ala Thr Arg Lys Ser Ala  
645 650

<210> 6

<211> 1314

<212> DNA

<213> artificial sequence

<220>

<223> fragment of the complete cDNA coding for  
the GBSSI of *Chlamydomonas reinhardtii*

<220>

<221> CDS

<222> (1)..(1314)

<223>

<400> 6

gcg ctg gac atc gtg atg gtt gct gct gag gtc gcc cct tgg tcc aag 48  
Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys  
1 5 10 15

acg ggc ggc ctg ggc gat gtg act ggt ggc ctg cct att gag ctg gtc 96  
Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val  
20 25 30

aag cgc ggc cac cgc gtc atg acc att gcc cct cgc tac gac cag tac 144  
Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr  
35 40 45

gct gac gcc tgg gac acc tcg gtg gtc gtg gac atc atg ggc gag aag 192  
Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys  
50 55 60

gtc cgc tac ttc cac tcc atc aag aag ggc gtg cac cgc gtg tgg att 240  
Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile  
65 70 75 80

gac cac ccc tgg ttc ctg gcc aag gtc tgg ggc aag acc ggc tcc aag 288  
Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys  
85 90 95

ctg tac ggc ccc cgc tcc ggc gct gac tac ctg gac aac cac aag cgc 336  
Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg  
100 105 110

ttc gcc ctg ttc tgc aag gcc gct att gag gct gcc cgc gtg ctg ccc 384  
Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro  
115 120 125

ttc ggc ccc ggc gag gac tgc gtc ttc gtg gcc aac gac tgg cac tcc Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser 130 135 140	432
gcc ctg gtg ccc gtc ctg ctg aag gac gag tac cag ccc aag ggc cag Ala Leu Val Pro Val Leu Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln 145 150 155 160	480
ttc acc aag gcc aag tcg gtg ctg gct atc cac aac atc gcc ttc cag Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln 165 170 175	528
ggc cgc atg tgg gag gag gct ttc aag gac acg aag ctg ccc cag gcc Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala 180 185 190	576
gcc ttt gac aag ctg gcc ttc tcg gac ggc tat gcc aag gtt tac act Ala Phe Asp Lys Leu Ala Phe Ser Asp Gly Tyr Ala Lys Val Tyr Thr 195 200 205	624
gag gcc acc ccc atg gag gag gac gag aag ccc ccg ctg acg gga aag Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys 210 215 220	672
acc tac aag aag atc aac tgg ctg aag ggt ggc att atc gcc gcc gac Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp 225 230 235 240	720
aag ctg gtg act gtg tcg ccc aac tac gcg acc gag atc gct gcc gat Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp 245 250 255	768
gcc gcc ggc ggt gtg gag ctg gac acc gtc atc cgc gcc aag ggc att Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile 260 265 270	816
gag ggc att gtg aac ggc atg gac att gag gag tgg aac ccc aag acc Glu Gly Ile Val Asn Gly Met Asp Ile Glu Glu Trp Asn Pro Lys Thr 275 280 285	864
gac aag ttc ctg tct gcg ccc tac gac cag aac agc gtc tac gcc ggc Asp Lys Phe Leu Ser Ala Pro Tyr Asp Gln Asn Ser Val Tyr Ala Gly 290 295 300	912
aag gcc gcc gcc aag gag gcc ctg cag gcc gag ctg ggc ctg cct gtg Lys Ala Ala Ala Lys Glu Ala Leu Gln Ala Glu Leu Gly Leu Pro Val 305 310 315 320	960
gac ccc acc gcc ccc ctg ttc gcc ttc atc ggc cgc ctg gag gag cag Asp Pro Thr Ala Pro Leu Phe Ala Phe Ile Gly Arg Leu Glu Glu Gln 325 330 335	1008
aag ggt gtg gac atc atc ctg gcc gcc ctg ccc aag atc ctg gcc acc Lys Gly Val Asp Ile Ile Leu Ala Ala Leu Pro Lys Ile Leu Ala Thr 340 345 350	1056
ccc aag gtg cag atc gcc atc ctg ggt acc ggc aag gcc gcc tac gag Pro Lys Val Gln Ile Ala Ile Leu Gly Thr Gly Lys Ala Ala Tyr Glu 355 360 365	1104

aag ctg gtg aac gcc atc ggc acc aag tac aag ggc cgc gcc aag ggc 1152  
 Lys Leu Val Asn Ala Ile Gly Thr Lys Tyr Lys Gly Arg Ala Lys Gly  
 370 375 380

gtg gtc aag ttc tcg gcg ccc ctg gcg cac atg ctc acc gcc ggc gcc 1200  
 Val Val Lys Phe Ser Ala Pro Leu Ala His Met Leu Thr Ala Gly Ala  
 385 390 395 400

gac ttc atg ctg gtg ccc tcg cgc ttc gag ccc tgc ggc ctg atc cag 1248  
 Asp Phe Met Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln  
 405 410 415

ctg cac gcc atg cac tac ggt acc gtg ccc gtg gta gcc tcc acc ggc 1296  
 Leu His Ala Met His Tyr Gly Thr Val Pro Val Val Ala Ser Thr Gly  
 420 425 430

ggc ctg gtc gac acc gtc 1314  
 Gly Leu Val Asp Thr Val  
 435

<210> 7  
 <211> 438  
 <212> PRT  
 <213> artificial sequence

<220>  
 <223> fragment of the complete cDNA coding  
 for the GBSSI of *Chlamydomonas reinhardtii*

<400> 7  
 Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys  
 1 5 10 15  
 Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val  
 20 25 30  
 Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr  
 35 40 45  
 Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys  
 50 55 60  
 Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile  
 65 70 75 80  
 Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys  
 85 90 95  
 Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg  
 100 105 110  
 Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro  
 115 120 125  
 Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser  
 130 135 140  
 Ala Leu Val Pro Val Leu Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln  
 145 150 155 160

Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln  
 165 170 175  
 Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala  
 180 185 190  
 Ala Phe Asp Lys Leu Ala Phe Ser Asp Gly Tyr Ala Lys Val Tyr Thr  
 195 200 205  
 Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys  
 210 215 220  
 Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp  
 225 230 235 240  
 Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp  
 245 250 255  
 Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile  
 260 265 270  
 Glu Gly Ile Val Asn Gly Met Asp Ile Glu Glu Trp Asn Pro Lys Thr  
 275 280 285  
 Asp Lys Phe Leu Ser Ala Pro Tyr Asp Gln Asn Ser Val Tyr Ala Gly  
 290 295 300  
 Lys Ala Ala Ala Lys Glu Ala Leu Gln Ala Glu Leu Gly Leu Pro Val  
 305 310 315 320  
 Asp Pro Thr Ala Pro Leu Phe Ala Phe Ile Gly Arg Leu Glu Glu Gln  
 325 330 335  
 Lys Gly Val Asp Ile Ile Leu Ala Ala Leu Pro Lys Ile Leu Ala Thr  
 340 345 350  
 Pro Lys Val Gln Ile Ala Ile Leu Gly Thr Gly Lys Ala Ala Tyr Glu  
 355 360 365  
 Lys Leu Val Asn Ala Ile Gly Thr Lys Tyr Lys Gly Arg Ala Lys Gly  
 370 375 380  
 Val Val Lys Phe Ser Ala Pro Leu Ala His Met Leu Thr Ala Gly Ala  
 385 390 395 400  
 Asp Phe Met Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln  
 405 410 415  
 Leu His Ala Met His Tyr Gly Thr Val Pro Val Val Ala Ser Thr Gly  
 420 425 430  
 Gly Leu Val Asp Thr Val  
 435

<210> 8

<211> 1593

<212> DNA

<213> artificial sequence



&lt;220&gt;

<223> fragment of the complete cDNA coding for  
the GBSSI of *Chlamydomonas reinhardtii*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1593)

&lt;223&gt;

&lt;400&gt; 8

gcg	ctg	gac	atc	gtg	atg	gtt	gct	gct	gag	gtc	gcc	cct	tgg	tcc	aag	48
Ala	Leu	Asp	Ile	Val	Met	Val	Ala	Ala	Glu	Val	Ala	Pro	Trp	Ser	Lys	
1				5					10					15		

acg	ggc	ggc	ctg	ggc	gat	gtg	act	ggt	ggc	ctg	cct	att	gag	ctg	gtc	96
Thr	Gly	Gly	Leu	Gly	Asp	Val	Thr	Gly	Gly	Leu	Pro	Ile	Glu	Leu	Val	
			20					25					30			

aag	cgc	ggc	cac	cgc	gtc	atg	acc	att	gcc	cct	cgc	tac	gac	cag	tac	144
Lys	Arg	Gly	His	Arg	Val	Met	Thr	Ile	Ala	Pro	Arg	Tyr	Asp	Gln	Tyr	
		35					40					45				

gct	gac	gcc	tgg	gac	acc	tcg	gtg	gtc	gtg	gac	atc	atg	ggc	gag	aag	192
Ala	Asp	Ala	Trp	Asp	Thr	Ser	Val	Val	Val	Asp	Ile	Met	Gly	Glu	Lys	
	50					55				60						

gtc	cgc	tac	ttc	cac	tcc	atc	aag	aag	ggc	gtg	cac	cgc	gtg	tgg	att	240
Val	Arg	Tyr	Phe	His	Ser	Ile	Lys	Lys	Gly	Val	His	Arg	Val	Trp	Ile	
65					70				75					80		

gac	cac	ccc	tgg	ttc	ctg	gcc	aag	gtc	tgg	ggc	aag	acc	ggc	tcc	aag	288
Asp	His	Pro	Trp	Phe	Leu	Ala	Lys	Val	Trp	Gly	Lys	Thr	Gly	Ser	Lys	
				85					90					95		

ctg	tac	ggc	ccc	cgc	tcc	ggc	gct	gac	tac	ctg	gac	aac	cac	aag	cgc	336
Leu	Tyr	Gly	Pro	Arg	Ser	Gly	Ala	Asp	Tyr	Leu	Asp	Asn	His	Lys	Arg	
			100					105					110			

ttc	gcc	ctg	ttc	tgc	aag	gcc	gct	att	gag	gct	gcc	cgc	gtg	ctg	ccc	384
Phe	Ala	Leu	Phe	Cys	Lys	Ala	Ala	Ile	Glu	Ala	Ala	Arg	Val	Leu	Pro	
		115					120					125				

ttc	ggc	ccc	ggc	gag	gac	tgc	gtc	ttc	gtg	gcc	aac	gac	tgg	cac	tcc	432
Phe	Gly	Pro	Gly	Glu	Asp	Cys	Val	Phe	Val	Ala	Asn	Asp	Trp	His	Ser	
	130					135					140					

gcc	ctg	gtg	ccc	gtc	ctg	ctg	aag	gac	gag	tac	cag	ccc	aag	ggc	cag	480
Ala	Leu	Val	Pro	Val	Leu	Leu	Lys	Asp	Glu	Tyr	Gln	Pro	Lys	Gly	Gln	
145					150					155					160	

ttc	acc	aag	gcc	aag	tcg	gtg	ctg	gct	atc	cac	aac	atc	gcc	ttc	cag	528
Phe	Thr	Lys	Ala	Lys	Ser	Val	Leu	Ala	Ile	His	Asn	Ile	Ala	Phe	Gln	
			165					170						175		

ggc	cgc	atg	tgg	gag	gag	gct	ttc	aag	gac	acg	aag	ctg	ccc	cag	gcc	576
Gly	Arg	Met	Trp	Glu	Glu	Ala	Phe	Lys	Asp	Thr	Lys	Leu	Pro	Gln	Ala	
		180						185					190			

gcc	ttt	gac	aag	ctg	gcc	ttc	tcg	gac	ggc	tat	gcc	aag	gtt	tac	act	624
Ala	Phe	Asp	Lys	Leu	Ala	Phe	Ser	Asp	Gly	Tyr	Ala	Lys	Val	Tyr	Thr	
	195						200						205			

gag gcc acc ccc atg gag gag gac gag aag ccc ccg ctg acg gga aag Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys 210 215 220	672
acc tac aag aag atc aac tgg ctg aag ggt ggc att atc gcc gcc gac Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp 225 230 235 240	720
aag ctg gtg act gtg tgg ccc aac tac gcg acc gag atc gct gcc gat Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp 245 250 255	768
gcc gcc ggc ggt gtg gag ctg gac acc gtc atc cgc gcc aag ggc att Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile 260 265 270	816
gag ggc att gtg aac ggc atg gac att gag gag tgg aac ccc aag acc Glu Gly Ile Val Asn Gly Met Asp Ile Glu Glu Trp Asn Pro Lys Thr 275 280 285	864
gac aag ttc ctg tct gcg ccc tac gac cag aac agc gtc tac gcc ggc Asp Lys Phe Leu Ser Ala Pro Tyr Asp Gln Asn Ser Val Tyr Ala Gly 290 295 300	912
aag gcc gcc gcc aag gag gcc ctg cag gcc gag ctg ggc ctg cct gtg Lys Ala Ala Ala Lys Glu Ala Leu Gln Ala Glu Leu Gly Leu Pro Val 305 310 315 320	960
gac ccc acc gcc ccc ctg ttc gcc ttc atc ggc cgc ctg gag gag cag Asp Pro Thr Ala Pro Leu Phe Ala Phe Ile Gly Arg Leu Glu Glu Gln 325 330 335	1008
aag ggt gtg gac atc atc ctg gcc gcc ctg ccc aag atc ctg gcc acc Lys Gly Val Asp Ile Ile Leu Ala Ala Leu Pro Lys Ile Leu Ala Thr 340 345 350	1056
ccc aag gtg cag atc gcc atc ctg ggt acc ggc aag gcc gcc tac gag Pro Lys Val Gln Ile Ala Ile Leu Gly Thr Gly Lys Ala Ala Tyr Glu 355 360 365	1104
aag ctg gtg aac gcc atc ggc acc aag tac aag ggc cgc gcc aag ggc Lys Leu Val Asn Ala Ile Gly Thr Lys Tyr Lys Gly Arg Ala Lys Gly 370 375 380	1152
gtg gtc aag ttc tgg gcg ccc ctg gcg cac atg ctc acc gcc ggc gcc Val Val Lys Phe Ser Ala Pro Leu Ala His Met Leu Thr Ala Gly Ala 385 390 395 400	1200
gac ttc atg ctg gtg ccc tgg cgc ttc gag ccc tgc ggc ctg atc cag Asp Phe Met Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln 405 410 415	1248
ctg cac gcc atg cac tac ggt acc gtg ccc gtg gta gcc tcc acc ggc Leu His Ala Met His Tyr Gly Thr Val Pro Val Val Ala Ser Thr Gly 420 425 430	1296
ggc ctg gtc gac acc gtc aag gag ggc gtc acc ggc ttc cac atg ggc Gly Leu Val Asp Thr Val Lys Glu Gly Val Thr Gly Phe His Met Gly 435 440 445	1344

gcc ctg aac ccc gac aag ctg gac gag gct gac gcc gac gcc ctg gcc 1392  
 Ala Leu Asn Pro Asp Lys Leu Asp Glu Ala Asp Ala Asp Ala Leu Ala  
 450 455 460

gcc acc gtg cgc cgt gcc agc gag gtg ttt gcg ggc ggc cgc tac ccc 1440  
 Ala Thr Val Arg Arg Ala Ser Glu Val Phe Ala Gly Gly Arg Tyr Pro  
 465 470 475 480

gag atg gtg gcc aac tgc atc agc cag gac ctg tcc tgg tcc aag ccc 1488  
 Glu Met Val Ala Asn Cys Ile Ser Gln Asp Leu Ser Trp Ser Lys Pro  
 485 490 495

gcc cag aag tgg gag ggc ctg ctg gag gag gtg gtg tac ggc aag ggc 1536  
 Ala Gln Lys Trp Glu Gly Leu Leu Glu Glu Val Val Tyr Gly Lys Gly  
 500 505 510

ggc gtg gcc acc gcc aag aag gag gag atc aag gtg ccc gtt gcc gag 1584  
 Gly Val Ala Thr Ala Lys Lys Glu Glu Ile Lys Val Pro Val Ala Glu  
 515 520 525

aag atc ccc 1593  
 Lys Ile Pro  
 530

<210> 9

<211> 531

<212> PRT

<213> artificial sequence

<220>

<223> fragment of the complete cDNA coding for  
 the GBSSI of *Chlamydomonas reinhardtii*

<400> 9

Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys  
 1 5 10 15

Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val  
 20 25 30

Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr  
 35 40 45

Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys  
 50 55 60

Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile  
 65 70 75 80

Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys  
 85 90 95

Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg  
 100 105 110

Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro  
 115 120 125

Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser  
 130 135 140

Ala	Leu	Val	Pro	Val	Leu	Leu	Lys	Asp	Glu	Tyr	Gln	Pro	Lys	Gly	Gln	
145					150					155					160	
Phe	Thr	Lys	Ala	Lys	Ser	Val	Leu	Ala	Ile	His	Asn	Ile	Ala	Phe	Gln	
				165					170					175		
Gly	Arg	Met	Trp	Glu	Glu	Ala	Phe	Lys	Asp	Thr	Lys	Leu	Pro	Gln	Ala	
			180					185					190			
Ala	Phe	Asp	Lys	Leu	Ala	Phe	Ser	Asp	Gly	Tyr	Ala	Lys	Val	Tyr	Thr	
		195					200					205				
Glu	Ala	Thr	Pro	Met	Glu	Glu	Asp	Glu	Lys	Pro	Pro	Leu	Thr	Gly	Lys	
	210					215					220					
Thr	Tyr	Lys	Lys	Ile	Asn	Trp	Leu	Lys	Gly	Gly	Ile	Ile	Ala	Ala	Asp	
225					230					235					240	
Lys	Leu	Val	Thr	Val	Ser	Pro	Asn	Tyr	Ala	Thr	Glu	Ile	Ala	Ala	Asp	
				245					250						255	
Ala	Ala	Gly	Gly	Val	Glu	Leu	Asp	Thr	Val	Ile	Arg	Ala	Lys	Gly	Ile	
			260					265					270			
Glu	Gly	Ile	Val	Asn	Gly	Met	Asp	Ile	Glu	Glu	Trp	Asn	Pro	Lys	Thr	
	275						280					285				
Asp	Lys	Phe	Leu	Ser	Ala	Pro	Tyr	Asp	Gln	Asn	Ser	Val	Tyr	Ala	Gly	
	290					295					300					
Lys	Ala	Ala	Ala	Lys	Glu	Ala	Leu	Gln	Ala	Glu	Leu	Gly	Leu	Pro	Val	
305					310					315					320	
Asp	Pro	Thr	Ala	Pro	Leu	Phe	Ala	Phe	Ile	Gly	Arg	Leu	Glu	Glu	Gln	
				325					330					335		
Lys	Gly	Val	Asp	Ile	Ile	Leu	Ala	Ala	Leu	Pro	Lys	Ile	Leu	Ala	Thr	
			340					345					350			
Pro	Lys	Val	Gln	Ile	Ala	Ile	Leu	Gly	Thr	Gly	Lys	Ala	Ala	Tyr	Glu	
		355					360					365				
Lys	Leu	Val	Asn	Ala	Ile	Gly	Thr	Lys	Tyr	Lys	Gly	Arg	Ala	Lys	Gly	
	370					375					380					
Val	Val	Lys	Phe	Ser	Ala	Pro	Leu	Ala	His	Met	Leu	Thr	Ala	Gly	Ala	
385					390					395					400	
Asp	Phe	Met	Leu	Val	Pro	Ser	Arg	Phe	Glu	Pro	Cys	Gly	Leu	Ile	Gln	
				405					410					415		
Leu	His	Ala	Met	His	Tyr	Gly	Thr	Val	Pro	Val	Val	Ala	Ser	Thr	Gly	
			420					425					430			
Gly	Leu	Val	Asp	Thr	Val	Lys	Glu	Gly	Val	Thr	Gly	Phe	His	Met	Gly	
	435						440					445				
Ala	Leu	Asn	Pro	Asp	Lys	Leu	Asp	Glu	Ala	Asp	Ala	Asp	Ala	Leu	Ala	
	450					455					460					

Ala Thr Val Arg Arg Ala Ser Glu Val Phe Ala Gly Gly Arg Tyr Pro  
 465 470 475 480

Glu Met Val Ala Asn Cys Ile Ser Gln Asp Leu Ser Trp Ser Lys Pro  
 485 490 495

Ala Gln Lys Trp Glu Gly Leu Leu Glu Glu Val Val Tyr Gly Lys Gly  
 500 505 510

Gly Val Ala Thr Ala Lys Lys Glu Glu Ile Lys Val Pro Val Ala Glu  
 515 520 525

Lys Ile Pro  
 530

<210> 10  
 <211> 23  
 <212> PRT  
 <213> artificial sequence

<220>  
 <223> fragment of GBSSI of Chlamydomonas reinhardtii

<400> 1  
 Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Gly Gly Lys  
 1 5 10 15

Thr Gly Gly Leu Gly Asp Val  
 20

<210> 2  
 <211> 23  
 <212> PRT  
 <213> artificial sequence

<220>  
 <223> fragment of GBSSI of Chlamydomonas reinhardtii

<400> 2  
 Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys  
 1 5 10 15

Thr Gly Gly Leu Gly Asp Val  
 20